Experiments designed to help the participants

Maximilian Kasy

7/10/2019

Why experiments?

- Are your programs effective in helping refugees?
  - How to find out?
- Possibility 1: Compare the outcomes of those who got the programs to others who didn't.
  - Problem: These groups might be different for other reasons.
- Think about a doctor prescribing a medical treatment.
  - Then the patients who got the treatment might die more often.
  - But only because they were more sick to begin with!
  - "Selection problem."

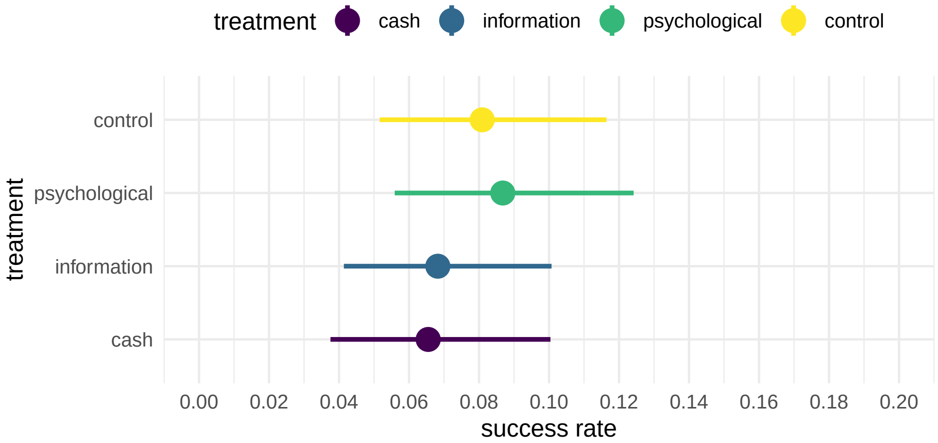
The standard way of doing experiments

- Possibility 2: Randomized experiment.
  - Create groups that are ex-ante similar,
    by randomly assigning participants to groups.
  - To compare apples with apples.
- Conventionally:
  - Divide the sample equally between treatments.
  - Wait until experiment is done.
  - Then compare average outcomes.
  - Use statistical tests to see whether there was any effect.

Drawbacks of conventional experiments

- This approach gets the causal effects right.
- And it gets precise estimates for every policy.
- But we need to wait a long time until we learn something.
- And we might not do the best we can for our participants.
- Think again of a medical experiment:
  - Suppose in the first few months, everybody who got the new treatment died.
  - Then you better stop the experiment!!!

Preliminary estimates for our experiment



- We already have suggestive evidence that the psychological treatment performs better.

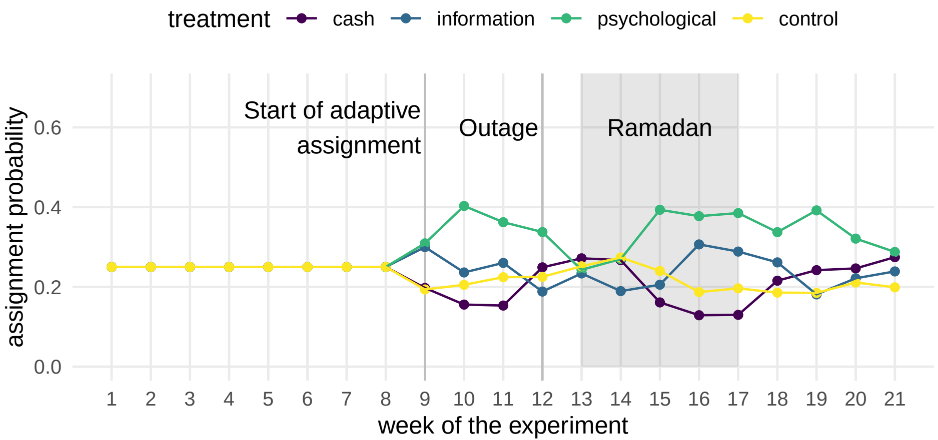
A different objective: Helping participants

- The standard approach is optimal when you want to get precise estimates of policy effects.
- But we want to instead help participants as much as possible.
- Cf. Immanuel Kant:
  "*Act in such a way that you treat humanity, whether in your own person or in the person of any other, never merely as a means to an end, but always at the same time as an end.*"
- This requires using the information we already have,
  when deciding which policy to assign people to.
- But we also want to continue learning,
  to do better in the future.

The exploitation / exploration tradeoff

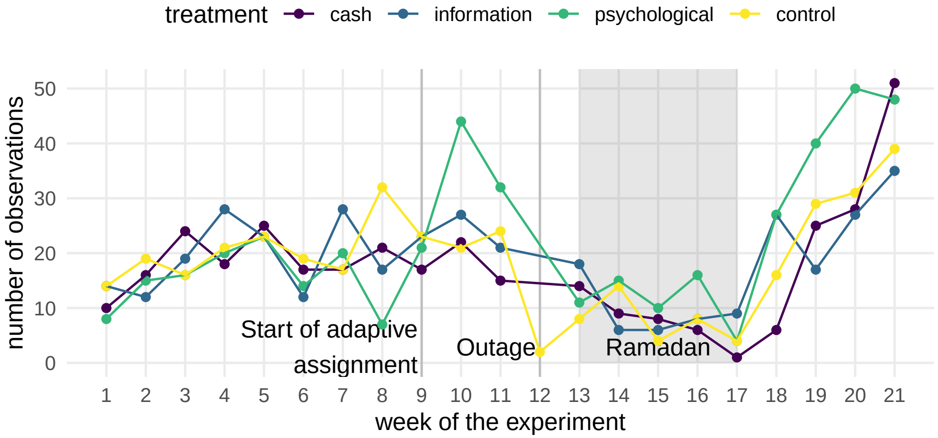
- Possibility 1: Assign each participant to the policy we currently think is best.
  - Good for the current participant.
  - Problem: We might stop learning,
    getting stuck with a sub-optimal policy.
- Possibility 2: Assign participants to each policy with fixed probability over time.
  - Good for learning policy effects.
  - But not optimal for current participants.
- Possibility 3: Optimal strategies shift to better performing policies over time.
- For instance *Thompson sampling*:
  - Assign each treatment with probability
  - equal to the current probability that it is optimal.

Assignment probabilities in our experiment



- As we learn that the psychological treatment does better, more participants are assigned to this treatment.

Assignment frequencies in our experiment



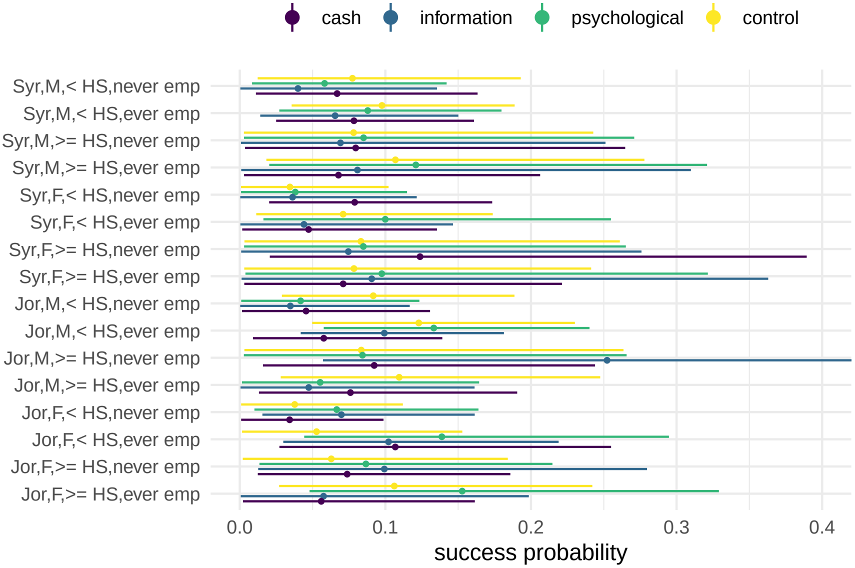
Targeting

- Not every policy is good for everybody.
- Some things work better
  - for those with more or less work experience,
  - for those with more ore less education,
  - for women or men.
- We can do better than just going with “one size fits all.”
- Try to get each group what works best for them.

Combining information

- Problem: For each group and policy, we might only have very few observations.
- This means averages are unreliable estimates.
- Solution: Combining information between groups.
- Estimate effect on a group by combining
  - their own average outcomes,
  - and the average outcomes for everybody else.
- *Bayesian hierarchical models* do this optimally.

Effect heterogeneity in our experiment



THANK YOU

For all your work in making this experiment happen!